# Fictional Objects within the Theory of Mental Files: Problems and Prospects

Zoltán Vecsey

A recent version of the mental file framework argues that the antirealist theory of fictional objects can be reconciled with the claim that fictional utterances involving character names express propositions that are true in the real world. This hybrid view rests on the following three claims: (i) character names lack referents but express a mode of presentation, (ii) fictional utterances introduce oblique contexts where character names refer to their modes of presentation, and (iii) modes of presentation are mental files. In this critical paper, I will argue that the proposed view runs into a number of theory-internal problems. These problems arise partly from the unclarities inherent in the notion of mental file, and partly from a mistaken semantics for character names. I will also argue that adherents of fictional realism can make use of the notion of mental file without encountering similar difficulties. | Keywords: *Antirealism, Fictional Objects, Mental Files, Character Names, Reference, Representation*

## 1 Fictional Objects in the Mental File Framework

The technical notion 'mental file' has been used recently by philosophers of language and theoretical linguists to explain the nature of singular thought and reference in natural language. Although there is no general consensus concerning the explanatory function of this term, it is widely agreed that the primary role of mental files is to store and manage information and, occasionally, misinformation about the objects we are somehow acquainted with. For example, Recanati (2012), a leading theorist of this approach, assumes that we can gain information/misinformation from a particular object when we stand in an epistemically rewarding relation to it. Sensory perception is the paradigm case of this kind of information gathering. We become aware of our immediate external environment by seeing or otherwise perceiving the sensory features of particulars. Acquaintance relations are usually interpreted normatively rather than logically or metaphysically in this area of research. We

open a mental file when there is an appropriate information channel between us and the object the file is about. This can be taken to be the normal or default situation. But it is not necessary (either logically or metaphysically) that there *actually* be such an information channel. Information gathering seems possible even in cases where acquaintance is merely imagined or simulated. Future-directed discourse is a good case in point. Our talk about the future is often based on acts of imagination. We have a natural inclination to make statements about future objects as if they were real existents in the world. In the broad sense of the term, we can thus specify objects that do not actually exist. We can attach proper names to "them", share our ideas about "them" and so forth. Imagination provides us with pieces of information that can be mentally stored in the usual manner. These and similar cases indicate that under certain circumstances mental files may be opened even in the absence of genuine epistemically rewarding relations.

This latter putative feature is what makes the notion of 'mental file' so attractive to theories of fictional objects. If we can indeed store and manage information/misinformation about purported objects in mental files without being actually acquainted with these purported objects, then by relying on this notion, we may try to give an account of how we can think of and talk about fictional persons and events.

Take for example the character of Sherlock Holmes in Conan Doyle's detective novel, A *Study in Scarlet*. Holmes is portrayed in the novel as being a detective. It is easy to check that Conan Doyle uses the proper name 'Sherlock Holmes' in his story consistently as if it were a genuinely referring singular expression. Yet we know that there is no such detective in the real world, external to the story. So regarding its semantic status, 'Sherlock Holmes' is a non-referring name. Given these two facts –that the main protagonist of Conan Doyle's novel is not a real person, and that 'Sherlock Holmes' does not refer to anything– it seems puzzling that we can gather so many pieces of information "about" Holmes and his deeds. We know very well that "he" is an outstanding detective, that "he" is a pipe-smoker, that "he" lives at 221B Baker Street, London, etc. Intuitively it seems we are able to think many singular thoughts involving these pieces of information. And it seems, again intuitively, that by expressing these thoughts we are able to make a potentially unlimited number of meaningful singular statements about the protagonist of the novel. How is this possible?

It is not easy to resolve this many-layered puzzle but it appears to be a good initial step to reflect on the way we collect information and misinformation about such fictional objects as Holmes. Adherents of the *antirealist* approach to fiction have recently offered an elegant and at first sight plausible explanation for this process.[1] The basic idea of this explanation is that, from the point of view of readers, fictional works should be conceived as prescriptions to *imagine*.[2] Novels and short stories prescribe us to imagine that things are

1 More precisely, antirealists are in agreement concerning the generic structure of the explanation, but they offer different versions of it. For example, see the works of Friend (2011, 2014) and Salis (2013).

2 The first occurrence of this idea is to be found in Walton (1990).

a certain way. In order to understand and appreciate a fictional work properly, we should follow as closely as possible the prescriptions originating from the narrative of that work. If the narrative tells us explicitly that there is a detective who smokes the pipe, lives at 221B Baker Street, London, etc., then we should imagine that there really is a detective who has exactly these properties. And if the narrative contains occurrences of 'Sherlock Holmes' in referring positions, then we should imagine that tokens of this name really refer to a person. Imagination does not require that we be committed to fictional objects; it requires merely that we be committed to pieces of information (and possibly misinformation) that can be extracted from the relevant narratives.

Given this basic idea, it is surely a well-motivated theoretical move to argue that this kind of information should be thought of as being collected in mental files. Antirealists who sympathize with the mental file conception of singular thought are obliged to say something about the nature of files. The common view, again originating from Recanati, is that a mental file usually consists of three components: the file itself with a certain label, the informational content of the file, and the reference-fixing relation that determines which object the file is about. As we have seen, when our targets are fictional works, the last component cannot be a genuine epistemic relation to an object external to a given narrative. But this does not generate a serious problem for the view. Thanks to our imaginative activities, labelled files can be opened and can be filled with pieces of information without the presence of external anchors. We can proceed broadly in the following manner. In reading the novel A Study in Scarlet, we encounter the character name 'Sherlock Holmes'. As a reaction to this reading experience, we open a mental file labelled with 'HOLMES'. All of the Holmes-relevant information that we can extract from the text of the novel will then be collected in the HOLMES file. We know, however, that our mental activities are governed in this process by the rules of imagination. And therefore we also know that when we deploy our HOLMES file we can refer only to an imagined person. According to the antirealist picture, this is why and how we can generate mental files on fictional characters and events with which we cannot, in principle, be acquainted.

Now the question arises whether the basic idea of this type of mental file theory is tenable or not. It is important to keep in mind that most adherents of the antirealist approach are convinced that *fictional objects do not exist*. On their view, there is simply no such fictional character as Holmes. But can such an allegedly nonexistent character be accounted for in terms of mental files?[3] In my own view, the short answer to this question is *no*. More cautiously, my claim is that we have good reasons to be skeptical concerning the explanatory power of the antirealists' mental file framework.

In what follows, I will focus my critical attention on the most recent version of the framework, elaborated and defended by Orlando (2017). Orlando's conception deserves attention for two reasons. First, the proposed framework is

[3] On this question, see also Murez and Smortchkova (2014).

sufficiently general for being a target of criticism. Second, Orlando supplements the standard conception of mental files with a semantic theory that gives a new twist to the ongoing debate about the interpretation of fictional statements. In Section 2, I briefly outline the main elements of Orlando's antirealist proposal. In Section 3, I try to point out that the proposed framework suffers from serious internal problems. Finally, in Section 4 I try to show that the notion of mental file is much less problematic when applied within the boundaries of a realist theorywhich acknowledges the existence of fictional objects.

## 2 Extending the Framework with a Two-Level Semantics

According to the mental file doctrine, if someone becomes acquainted with the novelist Jonathan Franzen, they open a file labelled with the mental name FRANZEN, and henceforth store or delete information/misinformation about Franzen exclusively in this very file. That is, they possess the individual file about Franzen, labelled with the mental name FRANZEN, and filled with descriptive concepts like 'born in Illinois', 'author of The Corrections', 'wearing spectacles', etc. Of course, different instances of the FRANZEN file may contain different sets of descriptive concepts. There might be readers who recognize Franzen as the 'author of The Corrections', others might know him as the 'author of Purity'. Differences in descriptive content do not affect the identity of the FRANZEN file, though. Competent readers will share the same file type because instances of this type ought to be individuated in the same way (i.e. by being related causally to Franzen) in every case.

Readers will be in a position to entertain *singular thoughts* about Franzen just in case they possess an instance of the FRANZEN file type. This is nearly self-evident. Yet it is not entirely obvious how mental files can be involved in expressing *singular propositions* about this person. The proposed explanation is that mental files should be thought of as devices of (mental) reference which are capable to refer to persons in roughly the same way as singular expressions refer in natural language. On this account, files are mental counterparts of proper names and, importantly, are supposed to be counterparts in the semantic sense of the word. If this is so, an utterance of the statement 'Franzen is the author of The Corrections' can express a mental or conceptual proposition about Franzen on the basis of the referential capacity of the FRANZEN file. Like its natural language counterpart, the expressed proposition counts as singular, since the FRANZEN file is grounded on causal relations to Franzen in roughly the same way as the proper name 'Franzen' is grounded causally on Franzen.

Orlando (2017, pp. 57–58) claims, on this basis, that mental files can be regarded as a constitutive component of the semantic content of singular utterances. This has already been recognized in the relevant literature. Recanati and many others have repeatedly argued that files play the role of non-descriptive Fregean modes of presentation. Seen from a semantic perspective, files as (non-descriptive) modes of presentation perform

a complex function: they are responsible for reference fixation, for cognitive significance, and coordination of information. Orlando proposes a two-level semantics where the content of singular expressions is constituted jointly by referents and such modes of presentation. It follows from this approach that an utterance of (1) has to be interpreted as expressing a two-level content composed by (1a) and (1b):

(1) Franzen is the author of The Corrections.

(1a) The *singular proposition* constituted by Franzen and the property of being the author of The Corrections.

(1b) The *conceptual proposition* constituted by the FRANZEN file and the descriptive concept ‘author of The Corrections’ contained in that file.

(1a) should be familiar, as it corresponds to the Russellian conception of singular propositions. This kind of content can be evaluated with respect to truth and falsity. If Franzen possesses the property of being the author of The Corrections, (1a) is true. At first glance (1b) may seem superfluous, since Russellian propositions are commonly assumed to express complete sentential contents without the intervention of modes of presentation. Orlando maintains, however, that (1b) does not determine (1a), contrary to what Fregeans might think. Rather, (1b) should be taken as representing an autonomous level of content. It is a mental or conceptual content that can be associated with the utterance of (1). And given that the FRANZEN file is a non-descriptive mode of presentation of Franzen, the conceptual proposition (1b) is not general but singular.

The mental file framework supplemented with the above two-level semantics can also be successfully applied to fiction – at least Orlando says so. The first important thing to note in this regard is that one can differentiate between three types of utterance in fictional narratives. As many have pointed out, there are *fictive*, *parafictive*, and *metafictive* utterances of sentences that differ sharply from each other with respect to their contextual background.[4]

Consider the following examples. The first token occurrence of the character name type ‘Sherlock Holmes’ in Conan Doyle’s oeuvre is to be found in his novel A Study in Scarlet, page 3, line 21:

(2) “You don’t know Sherlock Holmes yet”.

Since (2) is extracted from the text of the novel, it counts as a fictive utterance. Now compare (2) with (3):

(3) Sherlock Holmes was complemented perfectly by Dr. Watson.

(3) can be classified as a parafictive utterance.[5] One characteristic feature of

[4] The currently used terminology is not uniform. For example, instead of speaking of fictive, parafictive, and metafictive utterances, Thomasson (2003) uses the technical terms ‘fictionalizing discourse’, ‘internal discourse’, and ‘external discourse’.

[5] We can make a further distinction here between implicit parafictive utterances like (3) and explicit parafictive uttereances. The latter type uses prefixes such as ‘In work W’ or ‘According to the story S’. For present purposes, this distinction is irrelevant.

this type of utterance is that it is based on two narrative perspectives: (3) concerns the internal textual content of Conan Doyle's narrative but it paraphrases or restates this content from an external perspective. Metafictive utterances, in contrast, presuppose only a single perspective, a perspective that is external to the narrative.

(4) Sherlock Holmes is a fictional character.

In Conan Doyle's novel, Holmes is a detective, not a fictional character. But seen from the external perspective of literary criticism, Holmes is a fictional character. Accordingly, (4) counts as a paradigmatic metafictive utterance.

Intuitively, all of these utterances are meaningful and true either in the internal context of the novel *A Study in Scarlet* or outside of it. The observation that Sherlock Holmes and Dr. Watson are not real persons goes against this intuition, however. For if there are actually no such persons, then 'Sherlock Holmes' and 'Dr. Watson' are empty names, and thus utterances of such sentences as (2), (3), and (4) cannot express any proposition, which makes it hard to evaluate them as true.

This is a well-known problem that has been tackled by two main types of approach over the last decades. Realists argue that although Sherlock Holmes and Dr. Watson are not real persons, the Holmes character and the Dr. Watson character exist. Different brands of realism have elaborated different views on the nature of characters. Some conceive fictional characters as existing possibilia, others maintain that Holmes and Dr. Watson are created abstract objects and they are occasionally identified also with person-kinds existing in the same way as Platonic eternal idealities. What is common to all of these views is that they introduce an ontologically novel type of object for solving the above problem. If characters can be identified with possibilia, abstracta, or other types of objects, then character names can be taken to refer to these denizens of the world.[6] On this basis, realists can safely claim that utterances of (2), (3), and (4) express propositions, and are therefore true, as our intuition suggests.

In contrast,antirealists argue that fictional characters do not exist, and thus 'Sherlock Holmes', 'Dr. Watson' and other character names are empty. What justifies our intuition that utterances of (2), (3), and (4) are true is that by reading Conan Doyle's narrative readers imagine or assume that there are such persons as Holmes and Dr. Watson. This does not mean that utterances involving character names express singular propositions and are literally true. Such utterances are understood through an implicit paraphrase which typically takes the following form: according to an imaginative game authorized by the

[6] Fictive uses of character names still pose a problem for realists since tokens of 'Holmes' and 'Dr. Watson' do not refer to possibilia or abstracta or … in Conan Doyle's narrative. Therefore, realists usually argue that character names are empty in their fictive uses but parafictive and metafictive uses can refer back to characters that are already present at the primary textual level of the narrative. On this see, for example, Thomasson (2010).

novel *A Study in Scarlet*, such-and-such is the case.[7] Paraphrased in this way, fictional utterances express general propositions about the imaginative game rather than singular propositions about the characters of the narrative. What is said about Holmes and Dr. Watson is thus merely imaginatively true.

Orlando rightly observes that this situation confronts us with a dilemma. One option is that we interpret utterances like (2), (3), and (4) as expressing singular propositions about fictional characters at the cost of adopting a controversial ontology of objects. The other option is that we take (2), (3), and (4) to be parts of imaginative games at the cost of losing their capacity to express propositions about particular individuals. (see Orlando, 2017, p. 62)

Orlando's main contention is that the mental file framework sketched above enables us to avoid this dilemma. By adopting thisframework we can defend the antirealist theory of fictional characters and *at the same time* claim that utterances involving character names express singular propositions. The reasoning goes as follows. First, we should recognize that character names can be accounted for by the same two-level semantics as ordinary proper names. If 'Jonathan Franzen' has a referent (i.e. the person Franzen) and may be associated with a mode of presentation (i.e. the FRANZEN file), then 'Sherlock Holmes' should possess an identical or analogous set of semantic properties. The difference is, of course, that 'Sherlock Holmes' cannot be used to refer to the person Holmes since there is no such person. But then the supposed analogy between 'Franzen' and 'Holmes' disappears. Thus the second step in the reasoning consists of showing that the character name 'Holmes' should be taken to refer not to its customary referent (since there is no such thing) but to its mode of presentation (i.e. the HOLMES file). This amounts to showing that 'Holmes' is not empty even on the referential level of content. How can this be done? According to Orlando, utterances involving character names are not about the real external world. When readers talk about the protagonist of the novel *A Study in Scarlet*, they talk about something that has been created by Conan Doyle's fantasy. And it seems quite correct to assume that the products of the author's fantasy belong to the conceptual realm.

So when readers talk about the deeds and attributes of Holmes, they talk in fact about the conceptual content of the novel that was tokened first in Conan Doyle's mind. They can succeed in this only when they have a referential intention which is directed to this conceptual content. That is to say, by using the character name 'Holmes' readers of the novel must have the intention to refer obliquely to the mode of presentation of Holmes. The idea of oblique reference, of course, goes back to Frege, who once assumed that when referring expressions occur in the scope of an epistemic attitude verb like 'believe', theyrefer to their customary senses, not to their customary referents.[8]

[7] More precisely, this is a proposal which is characteristic of the Waltonian account of fiction. The main difficulty for this view is that metafictive utterances cannot be interpreted as belonging to authorized make-believe games. Antirealists claim, therefore, that utterances like 'Sherlock Holmes is a fictional character' should be seen as unauthorized make-believe games or betrayals of authorized make-believe games.

[8] Here, I would like to thank an anonymous reviewer for pointing out a flaw in an earlier draft of this paper.

In the case of a *fictive* utterance with a character name referential obliqueness means that a particular descriptive concept must be ascribed to a mental file. So 'Holmes' refers to the HOLMES file in (2) and the file is ascribed the descriptive concept 'is yet unknown to someone'. Since (2) is part of Conan Doyle's original novel, it can be taken to express this content automatically in the conceptual world of the narrative of that novel. Therefore, (2) comes out as true.

The *parafictive* utterance (3) requires a slightly different treatment because the descriptive concept 'was complemented perfectly by Dr. Watson' is not ascribed to the HOLMES file in the conceptual world of Conan Doyle's original narrative. If we want to find out whether or not (3) is true, we should analyse the narrative from an external perspective. It may turn out, after reading the novel, that Holmes and Dr. Watson have been portrayed by Conan Doyle as having a lot of complementary personal traits. If this is indeed the case, (3) accords with the conceptual content of the narrative. On this basis, (3) can also be judged as true.

The *metafictive* utterance (4) expresses a conceptual content that consists of the HOLMES file and the descriptive concept 'is a fictional character'. As in the case of (3), the descriptive concept is not part of the conceptual world of the original narrative. Moreover, 'is a fictional character' is a kind of content that is entirely incompatible with the internal perspective of the novel *A Study in Scarlet*. But approached from the external perspective of a reader who wants to talk about the ontological status of the character, (4) appears to be an ordinary, meaningful utterance. Orlando argues that this metafictive utterance can be accounted for by a hybrid interpretation. On the one hand, the utterer of (4) obliquely refers to the mode of presentation of Holmes. On the other hand, she ascribes a descriptive concept to the HOLMES file that does not accord with the conceptual content of the narrative. In other words, in this case, the referential shift is only partial: while the character name 'Holmes' changes its referent and refers to its mode of presentation, the predicate 'is a fictional character' retains its default semantic function and denotes the worldly property 'being a fictional character'. Thanks to this hybrid structure, (4) is partly about the conceptual world of the narrative and partly about the non-fictional world. That is why utterances of (4) can be interpreted as expressing true singular propositions.

## 3 Theory-Internal Problems

As we have seen above, Orlando's arguments are general enough to provide a satisfactory mental file framework for fiction. One noteworthy feature of the framework is that it can be applied to all types of fictional utterances. Unfortunately, despite its generality and applicability, the framework suffers from three systematic objections. Perhaps one of them may be reassuringly answered, but the other two seem to be troubling. Let's begin our survey with the weakest objection.

*Objection one: referential shift cannot be elicited by intention alone*. Orlando claims that a character name like 'Holmes' "seems not to refer to anybody." (Orlando, 2017, p. 66) I guess the "seems not to refer" here is only a polite way of saying that 'Holmes' is an empty singular expression. 'Holmes' lacks a referent because there is no real detective who lives at 221B Baker Street, London, smokes the pipe, etc. Of course, readers who participate in an authorised imaginative game use 'Holmes' as if it were an ordinary referring name. But it is important to note that imaginative games are unable to alter the semantic profile of names. If a name has been introduced into a narrative as an empty expression, it remains empty even if it is used within an imaginative game for referring to a person. 'Holmes' can be taken to refer to a detective in an imaginative game not because the imaginative game endows it with a referential capacity but because it is used in that game *as if* it were a referring name. In light of this, we can contend that referential emptiness is a constant semantic property of character names.

The mental file framework suggests otherwise. It is claimed that character names undergo a semantic shift and refer to mental files in all of their uses. After the shift has taken place, 'Holmes' ceases to be empty and starts to refer (obliquely) to the HOLMES file. The change in the semantic profile of the character name is supposed to be elicited by a specific sort of intention. Orlando calls this intention 'simulative'.

I find this picture rather implausible. My objection is not that this type of referential shift is in principle impossible. Indexicals and demonstratives refer via the intentions of speakers. So the content of an indexical expression or a demonstrative can be shifted by the referential intention of the speaker. This may happen even within an utterance of a single sentence. For illustration, consider a now-classic example of unbound pronouns from Kaplan's *Afterthoughts*: "You, you, you, and you can leave, but you stay." On Kaplan's view, it is the directing intention of the speaker that distinguishes between the referents of the token occurrences of 'you'. (see Kaplan, 1989, p. 589)[9] My objection is that the referents of character names cannot be shifted in this way. There is ample textual evidence that character names like 'Holmes' attempt to refer to persons. Although they do not succeed in this attempt, they are not sensitive to the changes of contextual factors like indexicals which have a two-dimensional (character/content) semantic structure. To repeat, readers of Conan Doyle's narrative may have a specific sort of intention to use'Holmes' for referring to a mental file or a mental representation but this will not yield the result that it in fact *refers* to a mental file or a mental representation.

One possible rejoinder to this objection is to point out that character names and other singular expressions are introduced into fictional narratives by simulative intentions. To adopt such a view would be tantamount to saying that character names refer to mental files from the very beginning of their

[9] It is worth noting that whether directing intentions are part of the semantic of demonstratives or belong to presemantics or pragmatics is a subject of debate.

career. Perhaps the first token occurrence of ‘Sherlock Holmes’ in the novel *A Study in Scarlet* refers already to a mental representation. Although this token occurrence seems to refer to a person within its host sentence, this is only a surface semantic effect.[10] Actually, Conan Doyle introduced the name of his protagonist to refer to its mode of presentation (i.e. the HOLMES file). So the argument may go. This would be a more plausible explanation for the alleged referential shift in the semantic profile of the name. If it is correct to assume that the profile of names depends, at least in part, on the semantically relevant aspects of their introduction, for example, in the semantic or communicative intentions of their introducers,thenit can be imagined that instances of a certain kind of name are designed so that they referto mentalobjects. The question is whether authors of fictional works introduce character names into their narratives in this manner. Regretfully, a definitive answer would require a lengthy excursion into the cognitive/psychological theory of artistic creation, which is beyond the scope of this paper. So let us leave this question open and turn instead to the second objection.

*Objection two: character names are supposed to perform two conflicting functions in fictional narratives*. As has already been mentioned, Orlando takes character names to refer to mental files. The character name ‘Holmes’ is supposed to refer to the HOLMES file, ‘Dr. Watson’ is supposed to refer to the DR WATSON file, and so forth. On hearing this, one may ask not only ‘what is the function of mental files?’, but also ‘what type of object are they?’. Orlando says that, from an ontological point of view, mental files are mental particulars. This does not clarify, however, whether they are concrete or abstract objects. Early advocates of the mental file theory like John Perry and Jerry Fodor have argued that files are objects in the mind or objects that are instantiated in the mind. These objects were conceived of as having causes and effects in the physical world. (see Fodor, 1990, pp. 23–25; Perry, 1980, p. 330) From this, it obviously follows that mental files were identified by these authors with concrete particulars. In a footnote, Orlando says that her own approach shares the ontological commitments of Fodor’s early work on mental representation. We may assume, then, that she would answer the question ‘what type of object are mental files?’ by saying that they are concrete particulars. If my reconstruction is correct, character names are supposed to refer ultimately to concrete objects in Orlando’s framework.

But this is, so to speak, only one aspect of the framework. The other aspect is that character names are supposed to refer to or stand for something abstract. The reason for this is the following. As already mentioned, according to Orlando, utterances of sentences like (2), (3), and (4) have to be interpreted as being about “something that has been created by an author’s

[10] A reviewer asks why we should take for granted that the first token of ‘Sherlock Holmes’ seems to refer to a person. According to the reviewer, this is counterintuitive since both the author and the readers know for a fact that there is no real person that is Sherlock Holmes. In response, I would say that it is better to keep the distinction between ‘seemings’ and ‘facts’: ‘Sherlock Holmes’ *seems* to refer to a person because it behaves in its first occurrence in Conan Doyle’s text as an ordinary personal name. What the author and the readers know about the existence/nonexistence of the character is, in my view, an independent issue.

imagination." (Orlando, 2017, p. 67) And the products of authorial imagination – fictional works and the characters portrayed in these works – must be seen as belonging to the realm of abstracta. This means that fictional works and fictional characters are not part of the physical world. Rather, they are identical with or part of the conceptual world that has been created by an author's artistic activity.

By applying this line of reasoning to the case of Holmes we get the following result. The main *character* of the novel *A Study in Scarlet* is the product of Conan Doyle's authorial imagination. The character is part of the conceptual world of the novel, from which it follows that it is an abstract object of some sort. Therefore, when the character is the subject matter of our utterances, token occurrences of 'Holmes' may be taken to refer to or stand for a certain abstract object.[11] In Orlando's own words: "our referential intention in using a fictional name can be construed as being oriented towards something not real (in the sense of belonging in the external world) but purely conceptual." (Orlando, 2017, p. 67)

The problem is that these two aspects of the framework are in conflict with each other. On the one hand, there is a semantic relation between the character name 'Holmes' and the HOLMES file. On the other, there is a semantic relation that relates 'Holmes' to the Holmes character. And this is something that cannot be integrated into a coherent semantic picture because 'Holmes' is related at once both to a concrete particular (i.e. HOLMES file) and to an abstract object. (i.e. the Holmes character)

*Objection three: the mental file framework is incompatible with the antirealist view of fictional objects*. There is a sharp disagreement between realists and antirealists on whether fictional objects exist. Realists believe that fictional objects are part of the overall inventory of what there is. This is not an innocent position because existing objects are typically thought of as being accessible through direct or indirect sensory experience; and it is fairly clear that Holmes, Dr. Watson, and their likes are not perceptible existents. Advocates of the realist view argue, therefore, that fictional objects are to be identified with a certain non-standard type of object. The most popular candidates are possibilia, created abstracta, and Platonic idealities. Fictional utterances are then interpreted as involving one of these types of non-standard objects. This ontological move saves the intuition that fictional utterances express singular propositions that can be either true or false. In contrast, antirealists are deeply convinced that our world does not contain any fictional objects. Fictional utterances appear to commit us to these objects but

[11] What complicates the picture is that Orlando favors an externalist conception of reference. On this conception, reference is *per definitionem* a relation to an "external" object. So it can be said that 'Holmes' *stands for* an abstractum, but it is incorrect to say that 'Holmes' *refers* to an abstractum. Not everyone shares this view. A well-known exception is to construe the relation of reference on the basis of negative free logic (Sainsbury 2005), which allows reference without referents. But there are also other alternatives. For example, Burge (2010) and Davies (2019) argue for a non-relational way of referring that can be successfully applied to abstract objects. Unfortunately, Orlando's framework does not take into consideration these developments.

from the antirealist's point of view, this is what it is: an appearance. We are willing to accept the existence of these objects because we entertain the utterances of fictional narratives by participating in authorised imaginative games. Accordingly, fictional utterances are to be taken to express true propositions about persons like Holmes or Dr. Watson only in an imaginative sense.

Orlando's mental file framework was designed to demonstrate that the antirealist theory of fictional objects can be reconciled with the claim that fictional utterances express propositions that are not imaginatively true, but instead true in the real world. But the framework cannot fulfill this promise.

There are at least two reasons for this. First, Orlando rejected the realist approach to fictional objects on the grounds that it embraces a non-standard ontology. One would expect, then, that her own approach is based on a standard ontological theory. Can mental files (i.e. concrete mental particulars) be incorporated into a standard classification scheme of objects? Although Orlando and other followers of Perry and Fodor find it self-evident that mental files constitute a natural kind, ontologists disagree with them on this point. The classificatory difficulty arises from the fact that mental objects of this type are "hybrid" existents, which satisfy the standard criteria both of concreteness and abstractness.[12] So it is not quite correct to suggest that the antirealist view of fiction can readily be paired with the mental file framework because both have equally parsimonious ontological commitments.

Second, and more importantly, it can be pointed out that the central claims of the mental file framework are incompatible with the antirealist view. While Walton (1990), Everett (2013), and other antirealists argue forcefully against the existence of fictional characters, Orlando seems to take an opposite view. She contends that if readers want to talk about the protagonist of a fictional work, then their referential intention is directed to *something* that belongs to the conceptual/abstract realm. And, on her view, this conceptual/abstract *something* exists contingently: it comes into being through an author's storytelling activity. But this is precisely what certain advocates of fictional realism claim. Artefactualists can happily accept that the protagonists of fictional works do indeed exist and that they can be classified as abstract objects.[13] Artefactualists can also agree with the claim that objects of fictional narratives like Holmes or Dr. Watson exist only contingently. Their abstract nature does not exclude that they are created objects. Many other products of our cultural activity come into being in a similar way: laws, institutions, marriages, etc., are paradigmatic abstract objects, but they do not and could not exist without the intervention of human intentional activity. This indicates rather clearly, I think, that Orlando's mental file framework is much closer to the artefactualists' position than it is to the antirealist view.

[12] For an overview of this issue, see McGinn (1980).

[13] The first systematic elaboration of the artefactualist position is to be found in Thomasson (1999). For a new version of the artefactualist view, see Vecsey (2019).

## 4 Mental Files from the Perspective of Fictional Realism

In the previous section, I pointed out that when we apply the mental file framework to the theory of fictional objects in a way similar to Orlando's, then the result will suffer from various theory-internal problems. A minor problem is that the framework assumes that the semantic profile of character names can be modified deliberately. Perhaps when authors introduce the names of their protagonists, they use these names from the outset as referring to something mental. Perhaps at least *some* of them use character names in this way. But this assumption needs empirical validation. Orlando's two-level semantics generates however a more serious problem. According to this view, character names are related both to mental files thought of as concrete particulars and to characters conceived of as being abstract objects. It is hard to see how this tension might be resolved within the proposed framework. It is also hard to see how the basic principles of the two-level semantics can be reconciled with the antirealist view which holds that there are no such things as fictional objects. The semantics has been so constructed that it allows for character names to be used referentially. Clearly, antirealists cannot tolerate this semantic claim, since it entails that character names do have referents and this means, ontologically speaking, that there are fictional objects.

In this last section, I will discuss briefly a possible way out of this theoretical impasse. As we have seen, fictional antirealists are in a difficult theoretical situation, because they have to reconcile two apparently incompatible theses. The first is a definitory claim. It says that it is a constitutive feature of mental files that they store and manage information/misinformation about objects. The second is the core ontological claim of the antirealist stance on fiction, which says that fictional objects do not belong to the overall inventory of what exists. The simplest and most often used antirealist strategy for reconciling these two claims is to adopt the Waltonian model of fiction and argue that fictional works should be conceived as prescriptions to imagine. Readers of fictional works have to imagine that things are in a certain way, for example, that Sherlock Holmes is a detective who lives at 221B Baker Street, London. It can be said, then, that although there are no fictional objects in reality, such objects feature in our imaginative acts and states. When readers collect and store information/misinformation about the protagonists of fictional works, their attention is directed to what should be imagined about these protagonists (i.e. the informational content of their own imaginatory acts and states). According to the resulting view, there is no obstacle to open mental files about merely imagined objects: even though Holmes lacks real existence, readers take it for granted, based on what they imagine, that "he" is an existing person.[14] In the end, files about spatiotemporal objects are supposed to differ from files about fictional objects only with respect to their type of reference. While reference is acquaintance-based in the first case, it is merely imagined in the second case.

[14] It is worth noting that Hansen and Rey (2016) sympathize with this view, but they argue for a mental file theory that is *neutral* with respect to whether the objects of the files are actual things or not.

Despite its *prima facie* plausibility, I do not consider the antirealist's explanatory strategy to be successful. The reason for skepticism is rooted in the very starting point of that strategy. In particular, it seems misleading to say that mental file theory in itself requires reconciling incompatible theses about fiction. I do not want to reject the definitory claim according to which mental files store and manage information/misinformation about objects. But I do think that the antirealist's ontological claim can be rejected, at least from the perspective of our everyday literary practices.

The question is whether we have access to the informational contents of literary works (i.e. literary texts) in the way antirealists assume. Is it really correct to say that imagination is our most direct and privileged epistemic relation to the textual level of works? I think the adequacy of the imagination-based conception is bounded by a more fundamental epistemic constraint: in order to imagine that a particular object o is so-and-so, we must already be acquainted with the text of a literary work which represents o linguistically as being so-and-so. This precondition may justly be thought of as a strong constraint because it narrows down the possible ways in which we may come to know of o to those that involve our language-based capacities. We simply need to read and process the relevant passages where o and its distinctive properties are portrayed. Imagination can only be activated after these passages have already been understood.

One might object that the epistemic priority argument does not undermine the antirealists' position because they may still argue that, even though our access to the contents of literary works must be mediated by language, fictional persons and events exist only in imagination. But the epistemic priority argument is more powerful than antirealists may think it is because it has an ontological consequence. If the text of a literary work represents o as being so-and-so, and our primary access to o is mediated by language, then our thought and talk about o should be taken as ontologically committing. By thinking and saying that Holmes is a detective we are committed to those representations, be they structurally simple or complex, that have the content or convey the information that *Holmes is a detective*.

A fictional realist *who follows this line of thought* may add that there is nothing more to being Holmes than being the content of these representations. The realist's general ontological point is that each fictional person can be identified with a specific set of representations. More precisely, the claim is that fictional persons have to be considered as embodiments of sets of interdependent and interconnected linguistic representations. 'Embodiment' stands here for a mental operation which binds separate but related representational elements into particular unities. We readers perform this operation rather easily when reading different passages of literary works. It does not take much reflection to recognize that scattered property descriptions like *is a detective*, *is a pipe-smoker*, or *lives at 221B Baker Street, London* belong to the same set of Holmes-representations in Conan Doyle's detective novel, *A Study in Scarlet*. The effective working of this operation is based, at least partly, on the fact that property descriptions of this type are capable of conveying informational

content. We understand perfectly well what it means to say that someone lives at 221B Baker Street, London. We understand this, even though the property description *lives at 221B Baker Street, London* lacks a language-external representatum.[15] Fictional realists may add, again, that this holds for informational contents in general: linguistic structures which are representation apt in literary works represent their target objects without being related to the language-external world.

In this regard, realists may rely on a conceptual distinction which was initially stated by Nelson Goodman (1968). In analysing the issues of pictorial representation, Goodman argued that 'represent' may occasionally be taken as an unbreakable one-place predicate. Many artistic pictures represent existing objects, Goodman says, but there are also pictures that do not represent anything. A picture of a unicorn is one of these cases. Yet to say this sounds a bit paradoxical. What could it mean that a picture does not represent anything and yet is a picture of a unicorn? If 'represent' has to be interpreted as a two-place predicate with an argument place for objects, then the paradox cannot be resolved. We ought to talk about a particular object and attribute properties to it when we want to talk about a representation. A way out is to recognise that a picture representing a unicorn is aunicorn-representing-picture, or, for short, a unicorn-picture, not a picture of or about a unicorn. This helps mitigate the paradoxical effects of the statement that although there are no unicorns, there are pictures that represent them. Although Goodman's main target was the problem of nonexistence in pictorial representations and he was obviously not a realist with respect to the ontological status of fictional creatures such as unicorns, his conceptual innovation seems to be easily transferable to the linguistic domain. The crucial point lies in the following distinction: non-fictional representations are normally *representations of objects*, where 'represent' should be interpreted as a two-place predicate; in contrast, fictional representations are *object-representations*, where 'represent' should be interpreted as a one-place predicate. Thus, while the former have a world-relational structure, the latter are thoroughly non-relational.

Coming back for the last time to Conan Doyle's main protagonist, it is essential to understand the order of explanation that is characteristic of the above-sketched approach. First, the epistemic priority argument states that our primary access to Holmes is mediated by language. We do not have, and cannot have, any language-independent knowledge about this fictional character: there are simply no exclusively perceptual means for recognising and identifying "him". Second, in reading Conan Doyle's novel we come to know that the text represents Holmes in a great variety of ways. More accurately, we come to know that the text contains a large number of property descriptions that have a common feature: all of these descriptions provide some partial information about one and the same protagonist. On that basis, we unify the descriptions under the label 'Sherlock Holmes', and then identify the character with this representational unity. In other words, we recognise that Holmes

[15] At the time the novel was written there was no such address as 221B at Baker Street.

embodies this representational unity. And third, relying on the Goodmanian distinction between the contrasting types of representation, weassociate with the character the semantic property of non-relationality. That is, we regard the unified and embodied property descriptions as providing us with a Holmes-representation, instead of a representation of Holmes. In this way, we can emphasize that in order to understand the mode of operation of the character name 'Holmes' there is no need to relationally refer to the facts and state of affairs of the language-external world.

The overall picture that arises from these short observations offer us three important lessons for the present context: (i) the imagination-first based approach to the problem of fictional objects is not mandatory; (ii) contrary to the antirealist doctrine, fictional objects do exist; and (iii) by applying the notion of non-relational representation, one can save the intuition concerning the informativity and understandability of literary texts. If this picture is correct, as I think it is, then realists can explain the possible connection between the theory of fictional objects and the mental file framework more easily than antirealists do.

Given that realists acknowledge the real, not only imagined existence of fictional objects, they can make use of the notion of mental file to provide an explication of how we store and manage information/misinformation about such objects, and this may be done in more than one way. They could argue, as above, that mental files have the function of binding together non-relational linguistic representations that readers gather from their reading experiences. (Vecsey, 2019) Alternatively, they may argue that fictional objects are created types and that readers refer to purported tokens of such types through mental files. (Terrone, 2017) Or they may claim, from the perspective of Discourse Representation Theory, that fictional objects are vicariously anchored entity representations that are stored in files. (Kamp, 2015) There are also other related options which take non-relationality as a property of purely intentional representations (Rey, 2003), or as a property of concepts. (Sainsbury, 2018) Which of these options is the most appropriate for the mental file framework depends, of course, on further details of the realist's view, but an in-depth discussion of this issue would require another paper.

## References


Burge, T. (2010) *Origins of objectivity*. First edition. Oxford; New York, NY: Oxford University Press.

Davies, J. E. (2019) 'Towards a Theory of Singular Thought about Abstract Mathematical Objects', *Synthese*, 196(10), pp. 4113–4136. doi: 10.1007/s11229-017-1644-0.

Everett, A. (2013) *The nonexistent*. First edition. Oxford; New York, NY: Oxford University Press.

Fodor, J. (1990) *A theory of content and other essays*. First edition. Cambridge, MA: Cambridge University Press.

Friend, S. (2011) 'The Great Beetle Debate', *Philosophical Studies*, 153, pp. 183–211. doi: 10.1007/s11098-009-9485-4.

Friend, S. (2014) 'Notions of Nothing', in García-Carpintero, M. and Marti, G. (eds) *Empty representations: Reference and non-existence.* Oxford; New York, NY: Oxford University Press, pp. 307–332.

Goodman, N. (1968) *Languages of art: an approach to a theory of symbols*. First edition. Indianapolis, Bobbs-Merrill.

Hansen, C. and Rey, G. (2016) 'Files and Singular Thoughts without Objects or Acquaintance: The Prospects of Recanati's (and Others') "Actualism"', *Review of Philosophy and Psychology*, 2, pp. 421–436. doi: 10.1007/s13164-015-0246-3.

Kamp, H. (2015) 'Using Proper Names as Intermediaries between Labelled Entity Representations', *Erkenntnis*, 80, pp. 263–312. doi: 10.1007/s10670-014-9701-2.

Kaplan, D. (1989) 'Afterthoughts', in Almog, J., Perry, J. and Wettstein, H. (eds.) *Themes from Kaplan*. Oxford; New York, NY: Oxford University Press, pp. 565–614.

McGinn, C. (1980) 'Philosophical Materialism', *Synthese*, 44, pp. 173–206.

Murez, M. and Smortchkova, J. (2014) 'Singular Thought: Object-Files, Person-Files, and the Sortal PERSON', *Topics in Cognitive Science*, 6, pp. 632–646. doi: 10.1111/tops.12110.

Orlando, E. (2017) 'Files for Fiction', *Acta Analytica*, 32, pp. 55–71. doi: 10.1007/s12136-016-0298-8.

Perry, J. (1980) 'A Problem about Continued Belief', *Philosophical Quarterly*, 61, pp. 317–332.

Recanati, F. (2012) *Mental files*. Oxford; New York, NY: Oxford University Press.

Rey, G. (2003) 'Intentional Content and a Chomskyan Linguistics', in Barber, A. (ed.) *Epistemology of language*. Oxford: Oxford University Press, pp. 140–186.

Sainsbury, M. (2005) *Reference without Referents*. First edition. Oxford; New York, NY: Oxford University Press.

Sainsbury, M. (2018) *Thinking about Things*. First edition. Oxford; New York, NY: Oxford University Press.

Salis, F. (2013) 'Fictional Names and the Problem of Intersubjective Identification', *Dialectica*, 67, pp. 283–301. doi: 10.1111/1746-8361.12031.

Terrone, E. (2017) 'On Fictional Characters as Types', *British Journal of Aesthetics*, 57, pp. 161–176. doi: 10.1093/aesthj/ayw091.

Thomasson, A. L. (1999) *Fiction and metaphysics*. First edition. Cambridge, MA: Cambridge University Press.

Thomasson, A. L. (2003) 'Speaking of Fictional Characters', *Dialectica*, 57, pp. 207–226. doi: 10.1111/j.1746-8361.2003.tb00266.x.

Thomasson, A. L. (2010) 'Fiction, Existence and Indeterminacy', in Woods, J. (ed.): *Fictions and models: New essays*. Munich: Philosophia Verlag, pp. 109–148.

Vecsey, Z. (2019) *Fiction and Representation*. First edition. Boston/Berlin: De Gruyter.

Walton, K. L. (1990) *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. First edition. Cambridge, MA: Cambridge University Press.

Zoltán Vecsey
MTA-DE-SZTE Research Group for Theoretical Linguistics
H-6722 Szeged, Egyetem utca 2.
Hungary
vecseyz@freemail.hu